

We Claim:

1. A method for language modelling of mixed language expressions comprising the steps of:

5 storing word equivalence probabilities relating to words of a first language and words in at least one other language;

generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities;

10 generating monolingual next word hypothesis probabilities in the first language based upon the monolingual word history; and

15 determining a probability of the next word in the mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities.
- 20 2. The method as claimed in claim 1, further comprising the step of summing the products of word equivalence probabilities with respective monolingual next word hypothesis probabilities.
3. The method as claimed in claim 1, wherein the monolingual next word hypothesis probability is a statistical language model .

25
4. The method as claimed in claim 1, further comprising the step of converting a mixed language word sequence to a monolingual word sequence using word equivalence probabilities.

30
5. The method as claimed in claim 1, further comprising the step of determining the word equivalence probabilities based upon a parallel text corpus that has

corresponding expressions in the first language and the at least one other language.

5 6. The method as claimed in claim 1, further comprising the step of determining a probability of a foreign language next word hypothesis given a base language word history.

10 7. The method as claimed in claim 1, further comprising the step of using a parallel text corpus that has corresponding expressions in the first language and the at least one other language.

8. A computer program product for language modelling of mixed language expressions, the computer program product comprising computer software recorded on a computer-readable medium for performing the steps of:

15 storing word equivalence probabilities relating to words of a first language and words in at least one other language;

20 generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities;

generating monolingual next word hypothesis probabilities in the first language based upon the monolingual word history; and

25 determining a probability of the next word in the mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities.

30 9. A computer system for language modelling of mixed language expressions, the computer system comprising computer software recorded on a computer-readable medium for performing steps of:

storing word equivalence probabilities relating to words of a first language and words in at least one other language;

5 generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities;

10 generating monolingual next word hypothesis probabilities in the first language based upon the monolingual word history; and

determining a probability of the next word in the mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities.